

ABSTRACT

The content of Web sites or other information within a markup format is automatically translated using an appropriate script written in the conversion language to “blindly” process a large number of Web sites. These implementations may employ an ECMAScript interpreter, a tier architecture, an SGML parser and dynamic tree-to-tree transformations. The tier architecture is used to control multiple target requests, grouping and organizing responses into markup documents. The SGML parser can provide fault-tolerant analysis of markup documents to make them conform to XML standards. The SGML parser can generate the tree of the resulting document as a dynamic mode representing the content of the original data. Dynamic tree-to-tree transformation is provided in general via a “template / match / select” script, and may also use such tools as an ECMAScript interpreter, a regular expression search, direct access to nodes by DOM navigation, and a transformation and service environment.